Available online at http://www.ijasrd.org/in

International Journal of Advanced Scientific Research & Development

Vol. 03, Iss. 02, Ver. II, Apr – Jun' 2016, pp. 149 – 154

e-ISSN: 2395-6089
p-ISSN: 2394-8906

DIFFERENT SENTIMENT ANALYSIS OF TEXT-TO-SPEECH SYNTHESIZE USING FUZZY NEURAL NETWORK FOR ENGLISH

B. Sudhakar

Assistant Professor, Department of Electrical Engineering, Annamalai University, Chidambaram, Tamil Nadu, India.

ARTICLE INFO

Article History:

Received: 20 Jun 2016;
Received in revised form: 29 Jun 2016;
Accepted: 29 Jun 2016;
Published online: 30 Jun 2016.

Key words:

Text to Speech Synthesizer (TTS), Fuzzy Neural Network (FNN), Neural Network (NN).

ABSTRACT

In urbanized countries, the customer service with automated system in speech synthesis has been the recent trend. A novel technique like Fuzzy Neural Network (FNN) based TTS system for the English language has been proposed for improve the naturalness of the automated speech synthesis systems. Different emotional sentences expressing “Happy”, “Fear”, “Neutral” and “Sad” have been given as input for this system. Depends on the emotions, it uses a group of fuzzy rules to segregate the sentences to recognize the respective emotions. The drawbacks of neural network has been eliminated by FNN incorporated in TTS system. It assigns specific labels for different emotions. Then the suitable emotion based speech output has been generated from the system. The genuineness of the FNN based TTS output is calculated through the comparative performance analysis with respect to the recorded human speech in the noise free environment. The spectral mismatch and amplitude variations of the resultant waves plays an vital task to measure the genuineness of the TTS output.

Copyright © 2016 IJASRD. This is an open access article distributed under the Creative Common Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

The TTS systems have been widely utilised in numerous real time applications for robotic speech synthesis systems. There has been a incredible growth and development in the area of Text to speech in most recent years by introducing many efficient methods to improve communication in efficient manner. Sentiment examination has become one of the important fields for research in the computational linguistics and is used for investigate the people's expression in either speech or text[1]. It is also called as opinion mining as it analyzes the people's opinions, attitudes and emotions.

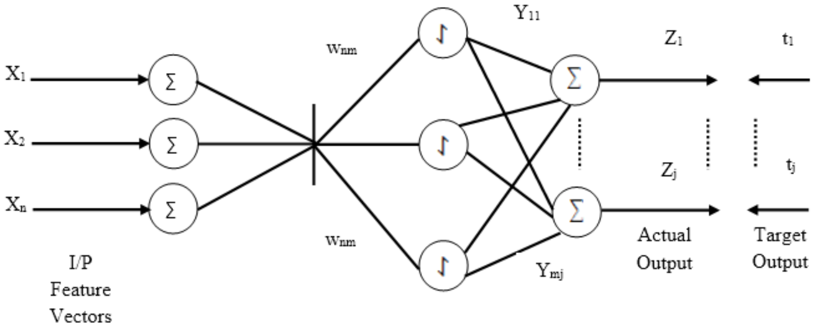
How to cite this article: Sudhakar, B., (2016). “Different Sentiment Analysis of Text-to-Speech Synthesize Using Fuzzy Neural Network for English”. *International Journal of Advanced Scientific Research & Development (IJASRD), 03* (02/II), pp. 149 – 154.

In modern years, emotion investigation has been widely used for examine the explanation and feedbacks in TTS synthesis. Usually, emotion investigation links under the group of categorization task. Emotion scrutiny steps process the input text through segregate positive, negative and neutral sentences[2][3]. TTS system has a number of real time community applications such as retail banking, railway stations, etc.

BACK PROPAGATION NEURAL NETWORK

ANN is a computational machine that implements an algorithm that is specified by the input nodes, output nodes, layers of hidden nodes between them, connecting lines, functions at the transforming nodes and weight multipliers $w_1,...,w_n$ on these lines, where, w is the weights of the vector[4][5]. The NN is a combination of several functions nominated by the wholel NN function as $z = f_{(x,w)}$. Fig.1 shows a conventional back propagation type of NN. The perpendicular bar to the right of the input nodes combine the n lines jointly into a bus to each of m nodes in the hidden layer. At the hidden nodes the weighted vector parts are added and put through a sigmoid soft threshold function to go to the output nodes to form the outputs $Z_{(j)}$. Each output node evaluate linear combination of (weighted) line values for z.

Figure – 1: *Architecture of Back Propagation Neural Network*

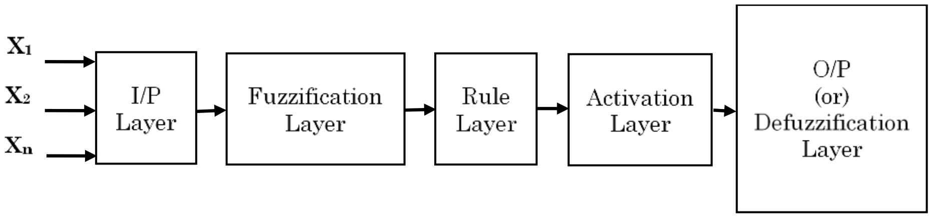


As the weights "w" varies, the NN function differs and so a exacte form of neural network is an unlimited relations of neural networks. A specified application must train "w" so that the NN outputs a near approximation "z" to the exact target $t_{(j)}$ for any input $x_{(j)}$. This is done by training on a set $\{(x_{(j)}, t_{(j)}) : j = 1,...,n\}$ of exemplar vector pairs of inputs and targets where the input vectors symbolize adequately well the piece of the input space from which the input vectors to be classified will be drawn. The trained NN then can interpolate on an input vector x and match its output "z" to the adjacent objective $t_{(q)}$ to offer the class of x.

FUZZY NEURO SYSTEM

Recently the fuzzy systems became prominent in engineering application The difficulty of discover membership functions and suitable policy is regularly a exhausting process of attempt and error. This brings to the thought of applying knowledge algorithms to the fuzzy systems[6][7]. An efficient learning algorithm adapted by the neural networks, were offered as an substitute to computerize or to maintain the improvement of tuning fuzzy systems. The first applications were in speech processing, image processing, data analysis, data classification and support to administrative problems etc.[8][9] Neural networks and fuzzy systems can be connected together to hold them advantages and to heal individual illness. NNs introduce evalution features of learning in the fuzzy systems and obtain from them the interpretation and clearness of systems illustration. Thus, the drawback of the fuzzy systems are eliminated by the capacities of the neural networks. These techniques are balancing which justifies use together.

Figure – 2: *Block Diagram of FNN*

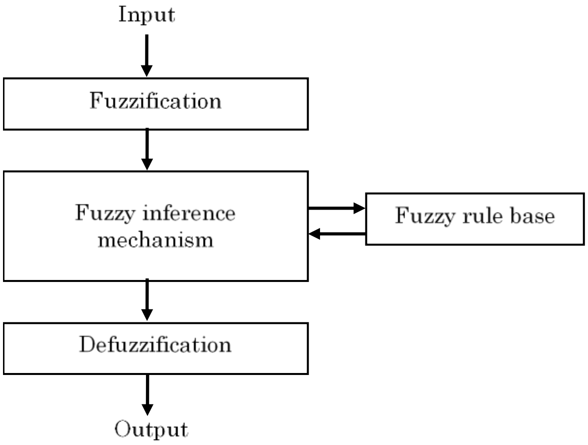


FNN uses a multi-layer perception network and a adapted back propagation training algorithm. FNN consist of five layers: input layer, fuzzification layer, ruling layer, activation layer, output (or) defuzification layer shown in Fig 2. The back-propagation algorithm is implemented for FNN training. The major characteristics of the FNN architecture are knowledge engineering capabilities[10][11]. The rules can be incorporated into FNN through some of the prefound membership functions and regulation and activation layer weights.

PROPOSED FNN BASED TTS SYSTEM RULE BASE

The proposed work focuses on TTS systems based on FNN to produce the output of different emotions for English language. The proposed framework is mainly focused on reviewing and gathering a group of character which is used for sentiment classification. The FNN classifies the input text depends on the sentiments derived from the sentence. The classifier import labels to each input text with suitable emotion depends on the effective attributes. It identifies the most appropriate emotion tag based on the features retrieved from the terms experiential in the text, which is frequently taken for a bag of words. FNN is utilised as a classifier to classify the text according to the emotions. FNN is the mixture of fuzzy theory and neural network. Since the both models are integrated, it could have better results. Both models are created depends on the functions of human brain, psychological reasoning and mental status of the humans.

Figure – 3: *Fuzzy Inference System*



An artificial neural network based on single neuron is utilised to defuzzify the fuzzy set as described in Fig 3. Each data in the fuzzy set are connected to the single neuron and it is transformed into the crisp outputs i.e., classified text with emotions. The single neurons accept the inputs from the fuzzy set and compute the bond value for each data using the transfer function. The output is the classification inference probability of each text. Table 1 shows the rule base of FNN. According to these rules, the words are categorized and the particular frequency is allotted for the words.

Table – 1: *Rule Base of FNN*

Emotions	**H**	**F**	**N**	**S**
H	PM	NME	NS	NM
F	NME	PM	NM	NS
N	NS	NM	PM	NME
S	NM	NS	NME	PM

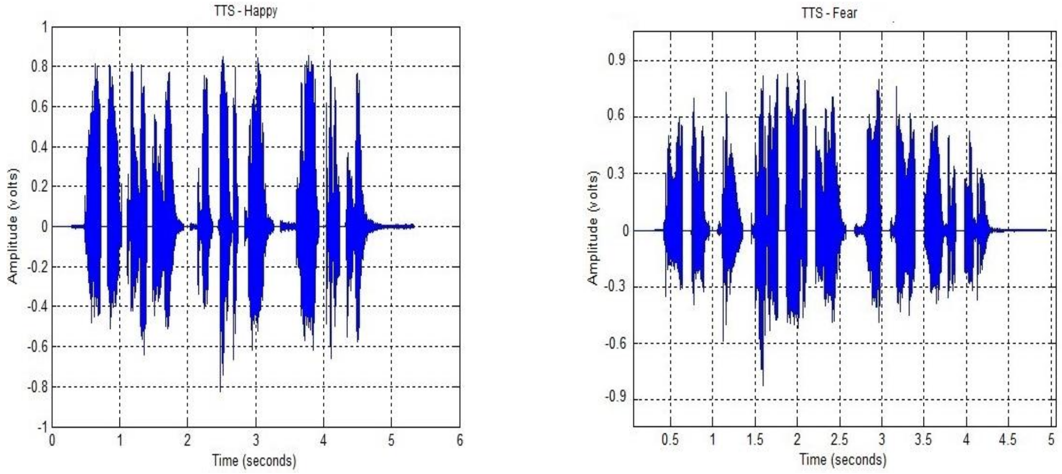
H: Happy; **F:** Fear; **N:** Neutral; **S:** Sad; **PME:** Positive medium; **PS:** Positive small; **PM:** Positive maximum; **NM:** Negative maximum; **NME:** Negative medium; **NS:** Negative small

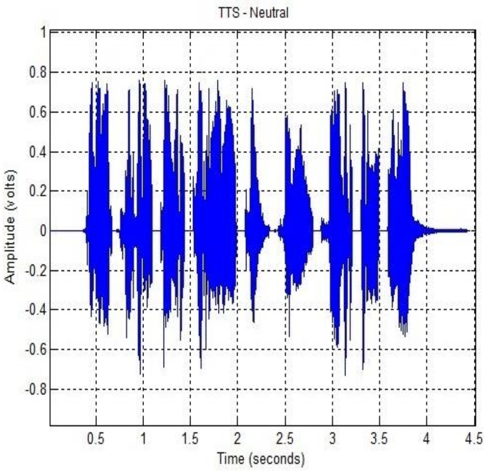
According to these rules, the words are categorized and the suitable frequency is allotted for the words. For example, if a positive word ‘H’ like ‘happy’ is given as input, the matching frequency is checked in the rule base Table 1 and the corresponding output ‘PM’ would be obtained which results in the corresponding expressive speech.

RESULTS AND DISCUSSION

The simulation results of “Happy”, “Fear”, “Neutral” and “Sad” emotions of the FNN based TTS system is shown in Fig 4 The recorded human voice samples have been taken from noise free environment. It is utilized as the reference for measure the performance of this TTS system.

Figure – 4: *Simulated TTS Output for Happy Fear Neutral Sad Voice using FNN*





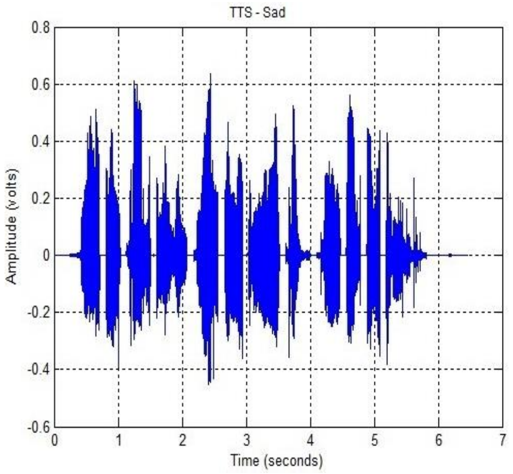


Table – 2: *Variations of Amplitude and Spectral Mismatch of FNN Based TTS Output for Different Emotions*

Emotions	Amplitude (in volts)	Spectral mismatch (in seconds)
Happy	0.9	0.065
Fear	0.79	0.093
Neutral	0.71	0.098
Sad	0.62	0.24

Table 2 shows the amplitude variations and spectral mismatch variations of the different emotional status for the proposed system. It is inferred that the naturalness has been measured through the amplitude variations and spectral mismatch variations of various emotional sentences with respect to recorded speech. When the amplitude increases and the spectral mismatch reduce the naturalness of the TTS system will be improved.

CONCLUSION

A novel emotion analysis using FNN is designed and implemented for TTS system of English Language. The required speech is generated by the concatenative speech synthesis approach with FNN. The given input text is analyzed and classified according to the emotions expressed in the sentence. The FNN has shown significantly better results with its training and testing capabilities. The emotions showing as “Happy”, “Fear”, “Neutral” and “Sad” are analyzed. The output speech clearly depicts the emotions of the user. The performance analysis has been carried out depends on the amplitude variations and spectral mismatch of the proposed system for different emotions. According to this analysis the “Happy” emotion produces the highest naturalness and “Sad” emotion produces lowest naturalness.

REFERENCES

[1] Kosko, B., (1992) *“Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence”.* Prentice Hall, Englewood Cliffs, NJ

[2] Lin, F. J., Hwang, W. J., & Wai, R. J., (1999) "A supervisory fuzzy neural network control system for tracking periodic inputs". *IEEE T. Fuzzy Syst., 7*(1): 41 – 52.

[3] Pang, B., Lee, L., & Vaithyanathan, S., (2002) *"Thumbs up? Sentiment classification using machine learning techniques".* Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP, 2002), pp: 79 – 86.

[4] Hu, M., & Bing, L., (2004) *"Mining and summarizing customer reviews".* Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD, 2004), pp: 168 – 177.

[5] Manning, C. D., Raghavan, P., & Schütze, H., (2008) *"Introduction to Information Retrieval".* Cambridge University Press, Cambridge, MA.

[6] Rudy, P., & Mike, T., (2009) "Sentiment Analysis: A Combined Approach". *J. Informetr., 3*(2): 143 – 157.

[7] Tiomkin, S., Malah, D., & Shechtman, S., (2010) "Statistical text-to-speech synthesis based on segment-wise representation with a norm constraint". *IEEE T. Audio Speech, 18*(5): 1077 – 1082.

[8] Schröder, M., Pirker, H., Lamolle, M., Burkhardt, F., Peter, C., & Zovato, E., (2011) *"Representing Emotions and Related States in Technological Systems".* In: Petta, P., R. Cowie and C. Pelachaud (Eds.), Emotion-Oriented Systems-the Humaine Handbook. Cognitive Technologies, Springer, Heidelberg, pp: 367 – 386.

[9] Drugman, T., Bozkurt, B., & Dutoit, T., (2012) "A comparative study of glottal source estimation techniques". *Computer Speech and Language, 26*(1): 20 – 34.

[10] Sudhakar B., & Bensraj, R., (2014) "An Efficient Sentence based Sentiment Analysis for Expressive Text-to-Speech using Fuzzy Neural Network". *Research Journal of Applied Sciences, Engineering and Technology, 8*(3): 378 – 386.

[11] Sangeetha, S., & Jothilakshmi, S., (2014) "Syllable based text to speech synthesis system using Auto associative Neural Network prosody prediction", *International Journal of Speech Technology, 17*(2): 91 – 98.

[12] Sudhakar, B., (2015). "Enhanced Sentimental Analysis for Speech Synthesis Based on Prosody Feature Modificatino Using TD-PSOLA Technique for English". *International Journal of Advanced Scientific Research & Development (IJASRD), 02* (02/I), pp. 59 – 64.